Generating and characterizing single- and multi-gene mutants 1 of the Rubisco small subunit family in Arabidopsis

Journal of Experimental Botany
doi:10.1093/jxb/eraa316 Advance Access Publication 6 July 2020
This paper is available online free of all access charges (see https://academic.oup.com/jxb/pages/openaccess for further details)

RESEARCH PAPER

Generating and characterizing single- and multigene mutants of the Rubisco small subunit family in Arabidopsis

Panupon Khumsupan[1], Marta A. Kozlowska[1], Douglas J. Orr[2,], Andreas I. Andreou[1], Naomi Nakayama[1,], Nicola Patron[3,], Elizabete Carmo-Silva[2,] and Alistair J. McCormick[1,*,]

[1] SynthSys & Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3BF, UK
[2] Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK
[3] Earlham Institute, Norwich Research Park, Norwich NR4 7UZ, UK

* Correspondence: alistair.mccormick@ed.ac.uk

Received 1 February 2020; Editorial decision 29 June 2020; Accepted 1 July 2020

Editor: Robert Sharwood, Australian National University, Australia

Abstract

The primary CO_2-fixing enzyme Rubisco limits the productivity of plants. The small subunit of Rubisco (SSU) can influence overall Rubisco levels and catalytic efficiency, and is now receiving increasing attention as a potential engineering target to improve the performance of Rubisco. However, SSUs are encoded by a family of nuclear *rbcS* genes in plants, which makes them challenging to engineer and study. Here we have used CRISPR/Cas9 [clustered regularly interspaced palindromic repeats (CRISPR)/CRISPR-associated protein 9] and T-DNA insertion lines to generate a suite of single and multiple gene knockout mutants for the four members of the *rbcS* family in Arabidopsis, including two novel mutants *2b3b* and *1a2b3b*. *1a2b3b* contained very low levels of Rubisco (~3% relative to the wild-type) and is the first example of a mutant with a homogenous Rubisco pool consisting of a single SSU isoform (1B). Growth under near-outdoor levels of light demonstrated Rubisco-limited growth phenotypes for several SSU mutants and the importance of the 1A and 3B isoforms. We also identified *1a1b* as a likely lethal mutation, suggesting a key contributory role for the least expressed 1B isoform during early development. The successful use of CRISPR/Cas here suggests that this is a viable approach for exploring the functional roles of SSU isoforms in plants.

Keywords: *Arabidopsis thaliana*, chloroplast, CRISPR/Cas9, photosynthesis, protoplasts, Rubisco, SpCas9, T-DNA.

Introduction

Rubisco (EC 4.1.1.39) catalyses the reaction between CO_2 and ribulose-1,5-bisphosphate (RuBP), and is responsible for net carbon gain in all oxygenic photosynthetic organisms including plants. Despite its importance, Rubisco is a relatively slow and error-prone enzyme that limits the efficiency of photosynthesis. This phenomenon is particularly evident in C_3 crop plants, where increasing the operating efficiency of Rubisco and reducing photorespiration are considered promising strategies for improving growth and productivity (Rae *et al.*, 2017; Kubis and Bar-Even, 2019; Simkin *et al.*, 2019; South *et al.*, 2019). In plants, Rubisco (Form IB) is composed of eight large subunits (LSUs) and eight small subunits (SSUs) that form an L_8S_8 complex (Bracher *et al.*, 2017). The LSU is encoded by a single, highly conserved gene on the chloroplast genome (*rbcL*), while SSUs are encoded by a family of nuclear *rbcS* genes that show significantly more diversity between species compared with *rbcL*. The size of the SSU family also differs between species, with up to 20 isoforms reported in cultivated polyploid wheat varieties (Sasanuma, 2001).

In Form I Rubisco, the eight LSUs form dimers that carry two active sites per dimer. The SSUs are located distal to the

active sites and are not required for carboxylation, but are essential to maximize Rubisco activity and are thought to provide structural stability to the L_8S_8 complex (Andersson and Backlund, 2008; van Lun *et al.*, 2011). A substantial body of work in the green alga *Chlamydomonas reinhardtii*, cyanobacteria, and higher plants has demonstrated that directed mutagenesis of the SSU, or expression of heterologous SSUs, can significantly modify the catalytic properties of Rubisco, including the turnover rate (k_{cat}), CO_2/O_2 specificity ($S_{c/o}$), and the ability to assemble the L_8S_8 complex (Schneider *et al.*, 1990; Paul *et al.*, 1991; Read and Tabita, 1992; Kostov *et al.*, 1997; Getzoff *et al.*, 1998; Spreitzer *et al.*, 2005; Genkov and Spreitzer, 2009; Ishikawa *et al.*, 2011; Esquivel *et al.*, 2013; Whitney *et al.*, 2015; Atkinson *et al.*, 2017; Fukayama *et al.*, 2019; Orr *et al.*, 2020). More recently, several native, specialized SSU isoforms have been shown to enhance the catalytic properties of Rubisco in plants (Morita *et al.*, 2014; Laterre *et al.*, 2017; Pottier *et al.*, 2018; Lin *et al.*, 2019, Preprint). This suggests that SSUs naturally play a role in modifying Rubisco performance and thus could be an important target for crop improvement.

The expression levels of SSUs also play a key role in regulating Rubisco abundance, and thus affect whole-plant nitrogen and carbon partitioning, and the overall capacity for carbon uptake (Rodermel, 1999). Previous work has shown that reducing SSU content by antisense RNA or T-DNA insertion results in a decrease in *rbcL* mRNA translation and, subsequently, a decrease in Rubisco content (Rodermel *et al.*, 1996; Wostrikoff and Stern, 2007; Izumi *et al.*, 2012; Ogawa *et al.*, 2012; Wostrikoff *et al.*, 2012; Atkinson *et al.*, 2017). The transcript abundances of SSU families have been explored in detail only for a small number of plant species, such as tomato, wheat, rice, and Arabidopsis (Wanner and Gruissem, 1991; Galili *et al.*, 1992; Suzuki *et al.*, 2009; Izumi *et al.*, 2012). Within a species, the strength of promoters for each SSU can vary significantly, resulting in a range of expression levels between different SSU isoforms. Furthermore, expression levels of individual SSUs can vary depending on tissue type and developmental stage, and in response to the growth environment (Wehmeyer *et al.*, 1990; Meier *et al.*, 1995; Ewing *et al.*, 1998; Day *et al.*, 2000; Morita *et al.*, 2014; Laterre *et al.*, 2017). Functional studies for individual SSUs in a species are challenging due to the high sequence homology between SSU isoforms within a family (Yamada *et al.*, 2019). Nevertheless, a better understanding of how the expression of different SSU isoforms is co-ordinated in response to the environment could lead to novel strategies to improve plant growth performances (Cavanagh and Kubien, 2014).

The model plant *Arabidopsis thaliana* (hereafter Arabidopsis) has four SSU genes, *rbcS1A*, *rbcS1B*, *rbcS2B*, and *rbcS3B* (hereafter *1A*, *1B*, *2B*, and *3B*, respectively), which are divided into A and B subfamilies based on linkage and sequence similarities (Krebbers *et al.*, 1988; Schwarte *et al.*, 2011). *1A* and *3B* are typically reported as the dominant SSU isoforms, while *1B* and *2B* are expressed at lower levels (Izumi *et al.*, 2012; Klepikova *et al.*, 2016). The Arabidopsis SSU family show signs of spatially overlapping and distinct expression during early leaf development (Sawchuk *et al.*, 2008). Although the response of Arabidopsis SSUs to environmental stimuli is relatively well studied (Dedonder *et al.*, 1993; Cheng *et al.*, 1998; Yoon *et al.*, 2001; Sawchuk *et al.*, 2008), a clear understanding of their impact on growth and performance is still lacking.

Gene knockout (KO) mutants generated by T-DNA insertion have been useful tools for functional studies in Arabidopsis (Izumi *et al.*, 2012; Atkinson *et al.*, 2017), but this approach does have limitations. First, T-DNA insertion sites are prone to small deletions, duplications, and filler sequence of unknown origin, while T-DNA lines have been shown to contain chromosomal translocations, and in some cases significant chromosomal rearrangements (Nacry *et al.*, 1998; Clark and Krysan, 2010). Thus, multiple T-DNA insertion lines for a given gene are typically required to verify experimental findings. However, in some cases, only a limited number of T-DNA lines are available for a given locus. Secondly, generating multiple gene KO lines is time consuming and not feasible for genes with loci in close proximity, as is the case for the three SSU B subfamily genes which, in Arabidopsis, are in a tandem array on chromosome 5 (Krebbers *et al.*, 1988; Niwa *et al.*, 1997).

To overcome this challenge, we have utilized a pooled clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated protein 9 (Cas9) approach and available T-DNA insertions lines to generate a novel suite of *rbcs* mutants. We produced new single *rbcs* mutants for all four *rbcS* genes, a double *rbcs* mutant (*2b3b*), and a triple *rbcs* mutant (*1a2b3b*). Molecular characterizations were performed to examine the impact of specific SSU mutations on protein and Rubisco contents, while physiological analysis under near-outdoor light levels (1000 μmol photons m^{-2} s^{-1}) provided novel insights into the contributions of different SSU isoforms to growth performance. This study serves as a proof of principle for future studies to examine the roles of different SSU isoforms in other species.

Materials and methods

Plant material and growth conditions

Arabidopsis [*Arabidopsis thaliana* (L.) Heyn. Col-0] seeds were sown on soil and stratified for 3 d at 4 °C, and grown at 22 °C, ambient CO_2, 70% relative humidity, and a photosynthetic photon flux (PPFD) of 200 μmol photons m^{-2} s^{-1} (standard lab conditions) or 1000 μmol photons m^{-2} s^{-1} (high light) supplied by cool white fluorescent lamps (Percival SE-41AR2, Clf Plantclimatics GmbH) in 12:12 h light:dark cycles. For comparison of different genotypes, plants were grown from seeds of the same age and storage history, harvested from plants grown in the same environmental conditions. Arabidopsis T-DNA insertion lines *1a* [GABI_608F01 (At1g67090)], *1b* [SAIL_755_D09 (At5g38430)], *2b* [GABI_324A03 (At5g38420)], and *3b* [SALK_117835 (At5g38410)] were sourced from the Nottingham Arabidopsis Stock Centre (NASC). The *1a2b* mutant (GABI_608F01; GABI_324A03) generated previously (Atkinson *et al.*, 2017) was backcrossed with a wild-type (WT) plant to remove potential background mutations. The *1a3b* mutant (GABI_608F01; SALK_117835) was provided by Hiroyuki Ishida, Department of Applied Plant Science, Tohoku University, Japan. Homozygous T_3 seed stocks for mutants generated via CRISPR/Cas9 in this study can be obtained through the NASC (http://arabidopsis.info) (NASC IDs N2109789 – N2109802).

Construction of CRISPR/Cas9 vectors

Plasmid vectors were assembled using the Plant MoClo Golden Gate modular cloning kit (Engler *et al.*, 2014). New Level 0 parts were made according to Patron *et al.* (2015). Level 0 vectors (100 ng each) carrying the UBI10 promoter, the SpCas9 coding sequence (Parry *et al.*, 2016), or the heat shock protein (HSP) terminator (Nagaya *et al.*, 2010) were assembled into the Level 1 Position 2 (L1P2) acceptor vector in a 20 µl assembly reaction [*Bsa*I (ThermoFisher Scientific, UK) (10 U), 1× Buffer G, T4 DNA ligase (ThermoFisher Scientific) (400 U), and 20 nmol ATP] as in Vasudevan *et al.* (2019). PCR amplicons of each complete guide RNA (gRNA; the spacer fused to the RNA scaffold) were combined with a Level 0 vector carrying the U6 promoter for assembly (see Supplementary Table S1 at *JXB* online). Each pair of gRNA expression cassettes were constructed in L1P3 and L1P4, respectively, as described in Raitskin *et al.* (2019). Four Level 1 transcriptional units [the pFAST selection marker (Shimada *et al.*, 2010) in L1P1, L1P2, L1P3, and L1P4] were then assembled into a Level 2 acceptor vector in a 20 µl assembly reaction [*Bpi*I (ThermoFisher Scientific) (10 U), 1× Buffer G (ThermoFisher Scientific), T4 DNA ligase (ThermoFisher Scientific) (400 U), and 20 nmol of ATP] (see Supplementary Data S1) as in Vasudevan *et al.* (2019).

DNA and RNA extraction, PCR, and RT–qPCR

DNA was extracted from a mature leaf as described in Li and Chory (1998). PCRs were performed as in McCormick and Kruger (2015) using gene-specific primers (Supplementary Table S2). Total RNA was isolated from leaves using the RNeasy plant mini kit (Qiagen, USA). Isolated RNA was treated with DNase (Qiagen) and reverse transcribed with random primers (Promega, USA). Gene-specific primers amplifying the unique 3' region of the transcript were used for quantitative reverse transcription–PCR (RT–qPCR) (Izumi *et al.*, 2012). A DNA fragment containing regions matching the target loci of the *rbcS* for RT–qPCR primers was synthesized (Gblock, IDT) (Supplementary Fig. S1). RT–qPCR calibration curves were constructed using known concentrations of the standard to quantify mRNA levels for each *rbcS* transcript pool. For quantitative analysis, an aliquot of cDNA derived from 4 ng of RNA was used (total volume 20 µl) with SYBR Green Master Mix (Eurogentec, Belgium).

CRISPR/Cas9 cassettes in protoplasts

Leaves from 4-week-old plants were cut vertically into 1 mm strips and digested in 10 ml of maceration glycine glucose (MGG) digestion solution as in Chupeau *et al.* (2013) containing cellulase 'Onozuka' R-10 [1.5% (w/v)] and Macerozyme R-10 [0.4% (w/v); Yakult Pharmaceutical, Japan] for 3 h. Released protoplasts were filtered from the digestate using a 70 µm cell strainer and washed three times with MGG not containing enzymes to remove traces of the enzyme solution and cell debris. Protoplasts were resuspended in MMM solution [0.4 M mannitol, 15 mM $MgCl_2$, 0.1% (w/v) MES (pH 8)] to a concentration of 5×10^5 cells ml^{-1} in a 5 ml glass test tube. For protoplast transformations, 8 µl of DNA (4 µg total) was added to 75 µl of the protoplast suspension, followed by addition of 83 µl of polyethylene glycol (PEG) solution [0.4 M mannitol, 0.1 M $Ca(NO_3)_2{\cdot}4H_2O$, 40% (w/v) PEG 4000 (pH 8)]. Following a 1 min incubation, 2 ml of MGG solution was added. Following a further 1 h incubation at room temperature, the protoplasts were centrifuged at 70 *g* for 5 min and the supernatant removed. Fresh MGG solution was added (100 µl), and transfected protoplasts were incubated in the dark for 18 h at room temperature. The target loci of each CRISPR/Cas9 vector were analysed by PCR of protoplast DNA extracts (Supplementary Table S2).

Expression of Cas9 and gRNA in Arabidopsis

Binary vectors (Level 2) were transformed into *Agrobacterium tumefaciens* (AGL1) for stable insertion in Arabidopsis by floral dipping (Clough and Bent, 1999). T_1 plants were screened for the presence of the transgene by the pFAST selectable marker (Shimada *et al.*, 2010), and for the presence of CRISPR/Cas9-mediated mutations by PCR and Sanger sequencing (Supplementary Table S2). Stable mutations in transgene-free T_2 plants were confirmed by Sanger sequencing.

Protein quantification and Rubisco content

Leaf samples (20–40 mg) were collected from 35-day-old plants, snap-frozen, and stored at –80 °C prior to extraction. Samples were ground rapidly in an ice-cold mortar and pestle in 200 µl of protein extraction buffer [50 mM Bicine-NaOH pH 8.2, 20 mM $MgCl_2$, 1 mM EDTA, 2 mM benzamidine, 5 mM ε-aminocaproic acid, 50 mM 2-mercaptoethanol, 10 mM DTT, 1% (v/v) protease inhibitor cocktail (Sigma-Aldrich, USA), and 1 mM phenylmethylsulfonyl fluoride] for ~1 min followed by centrifugation at 14 700 *g* at 4 °C for 1 min. Supernatant (90 µl) was then mixed with 100 µl of carboxyarabinitol-1,5-bisphosphate (CABP) binding buffer {100 mM Bicine-NaOH (pH 8.2), 20 mM $MgCl_2$, 20 mM $NaHCO_3$, 1.2 mM (37 kBq μmol^{-1}) [^{14}C]CABP}, incubated at room temperature for 25 min, and Rubisco content was determined via [^{14}C]CABP binding (Sharwood *et al.*, 2016). Bradford assay was used to determine total soluble protein in the same supernatant as prepared for Rubisco content analysis (Bradford, 1976).

Extracts were subjected to SDS–PAGE on a 4–12% (w/v) polyacrylamide gel (Bolt® Bis-Tris Plus Gel) (ThermoFisher Scientific, UK), transferred to a PVDF membrane, then probed with rabbit serum raised against wheat Rubisco at 1:10 000 dilution (Howe *et al.*, 1982) followed by LI-COR IRDye® 800CW goat anti-rabbit IgG (LI-COR Biosciences, USA) at 1:10 000 dilution, then viewed on an LI-COR Odyssey CLx Imager. The relative abundances of LSUs and SSUs were estimated densitrometically using Image Studio Lite (LI-COR Biosciences) and the values were means ±SE based on three immunoblots as in Atkinson *et al.* (2017)

Chlorophyll quantification

Leaf discs (20 mm^2 in total) were frozen in liquid N_2, powdered, and then mixed with 1 ml of ice-cold 80% (v/v) acetone, 10 mM Tris–HCl. Following centrifugation at 17 200 *g* for 10 min, chlorophyll was quantified according to Porra *et al.* (1989).

Measurement of photosynthetic parameters

Gas exchange and chlorophyll fluorescence were determined using a LI-COR LI-6400 portable infra-red gas analyser (LI-COR Biosciences) with a LI6400-40 leaf chamber (2 cm^2 area) on either the sixth or seventh leaf of 35- to 45-day-old rosettes grown under 200 µmol photons m^{-2} s^{-1} in large pots to generate leaf area sufficient for gas exchange measurements (Atkinson *et al.*, 2017). For all gas exchange experiments, leaf temperature and chamber relative humidity were 25 °C and ~65%, respectively. The response of the net CO_2 assimilation (A) to the intercellular CO_2 concentration (C_i) was measured at various CO_2 concentrations (50, 100, 150, 200, 250, 300, 350, 400, 500, 700, 900, and 1200 µmol mol^{-1}) under saturating light (1800 µmol photons m^{-2} s^{-1}) (Supplementary Fig. S2). Gas exchange data were corrected for CO_2 diffusion from the measuring chamber as in Bellasio *et al.* (2016). To calculate the maximum rate of Rubisco carboxylation (V_{cmax}), the A/C_i data were fitted to the C_3 photosynthesis model as in Ethier and Livingston (2004) using the catalytic parameters K_c^{air} and affinity for O_2 (K_o) values for WT Arabidopsis Rubisco at 25 °C from Atkinson *et al.* (2017). Estimates of the light- and CO_2-saturated photosynthetic electron transport rate (J_{max}) were not included as several of the SSU mutants were probably Rubisco limited even at high CO_2 concentrations. Maximum quantum yield of PSII (F_v/F_m) was measured using a Hansatech Handy PEA continuous excitation chlorophyll fluorimeter (Hansatech Instruments, UK) (Maxwell and Johnson, 2000).

Rosette area and biomass

Rosettes were imaged daily during growth experiments. Rosette area was calculated using iDiel Plant software (Dobrescu *et al.*, 2017). For biomass measurements, aerial rosette tissue was removed and weighed to

determine fresh weight. Samples were then dried in an oven (80 °C for 3 d) and weighed to determine dry weight.

Statistical analysis

Significant differences between sample groups were assessed by one-way ANOVA followed by Tukey's honest significant difference (HSD) post-hoc test (IBM SPSS Statistics Ver. 26.0, USA) for individual parameters. Difference in growth performance (as measured by rosette area) were assessed by repeated measures ANOVA followed by Tukey's HSD post-hoc test.

Results

Identification of T-DNA insertion lines and double knockout mutants

We initially performed a search for available Arabidopsis mutant lines on T-DNA Express (http://signal.salk.edu/cgi-bin/tdnaexpress) that carry a single homozygous T-DNA insertion, ideally located in exonic regions for members of the SSU gene family. We identified only one such line for *1A* (GABI_608F01) and *2B* (GABI_324A03) (Fig. 1A). The only available homozygous mutants for *1B* (SAIL_755_D09) and *3B* (SALK 117835) had insertion sites located in the 5'-untranslated region (UTR). For the latter, *3B* expression is reportedly reduced but not absent (Izumi *et al.*, 2012).

Double mutants *1a2b* and *1a3b* were generated previously by crossing available T-DNA lines (Izumi *et al.*, 2012; Atkinson *et al.*, 2017). Upon further characterization of *1a2b*, we observed that the mutant had a defective silique phenotype, with a reduced silique size and seed count per silique (Supplementary Fig. S3A, B). Backcrossing with a WT plant and then re-segregating *1a2b* in the F_2 generation removed the observed silique phenotype, indicating that this trait was not

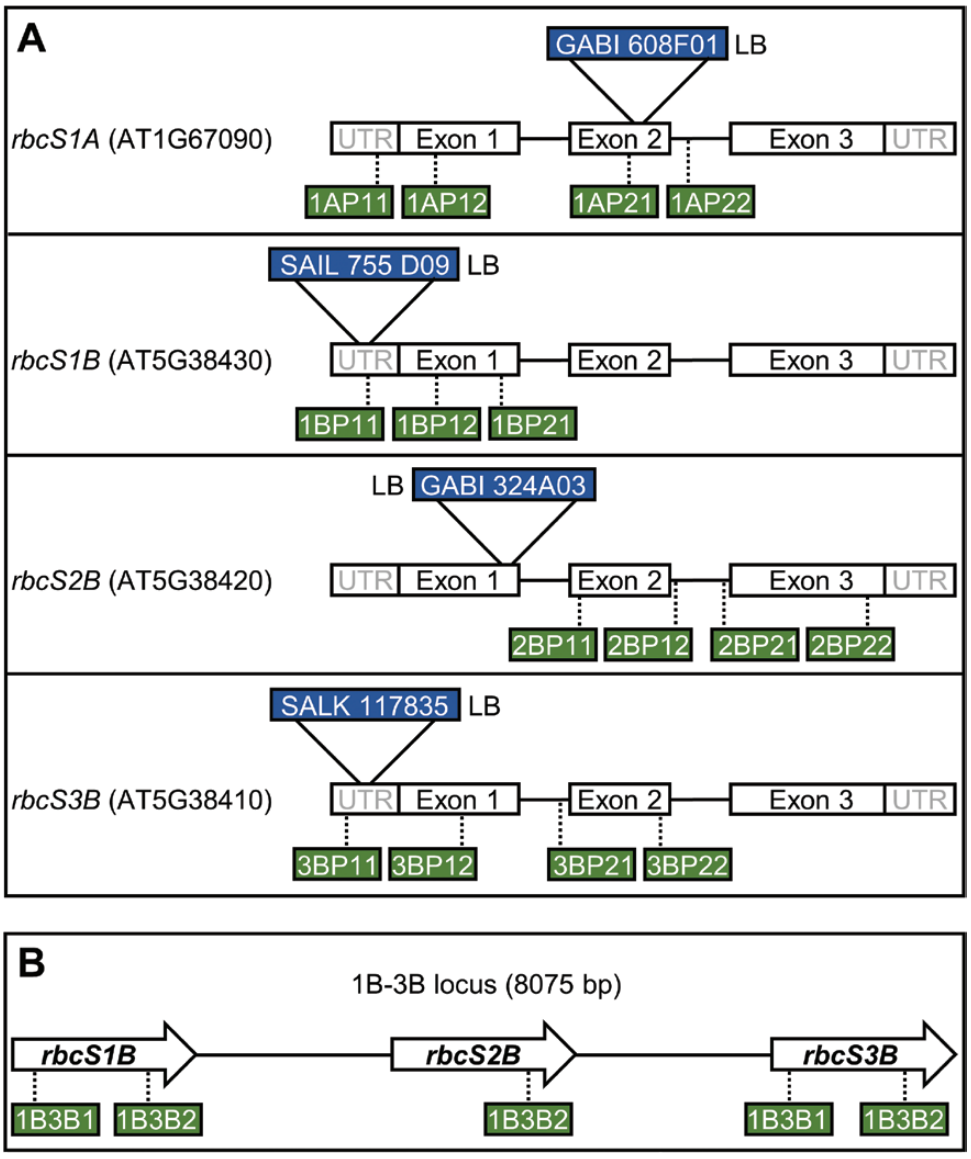

Fig. 1. The Rubisco small subunit gene family in *Arabidopsis thaliana*. (A) The sites of the T-DNA insertions for mutant lines used in this study are shown in blue, with the orientation of the left border (LB) indicated. The locations targeted by CRISPR/Cas9 are shown, with the names of the targeting gRNAs in green boxes (see Supplementary Table S1 and vector maps in Supplementary Data S1 for gRNA sequences). The 5'- and 3'-untranslated regions (UTRs) are shown for each gene. (B) The B subfamily is located in tandem in the *1B–3B* locus. A promiscuous pair of gRNAs targeting the three small subunits in the *1B–3B* locus is indicated.

attributable to the absence of 1A and 2B. Furthermore, the new *1a2b* showed a more robust growth phenotype compared with the original double mutant (Supplementary Fig. S3C). Single *rbcs* mutant lines for *1a* or *2b* did not show a silique phenotype. We identified plants in the segregating F_2 population with the silique phenotype or reduced growth, but with no T-DNA insertions in either *1A* or *2B*, indicating that these traits were possibly linked to a recessive heterozygous mutation in either the *1a* or *2b* parental line. The new *1a2b* double mutant was used for the remainder of this study. We also attempted to generate a *1a1b* double mutant by crossing the T-DNA insertion mutants for *1a* and *1b*, and were successful in generating a heterozygote F_1 (1A*1a*1B*1b*) line (Supplementary Fig. S4). We were not able to recover a *1a1b* mutant after screening 125 F_2 plants, but did identify a 1A*1a1b1b* line. Analysis of 15 F_3 progeny of the latter line failed to show a Mendelian distribution for the *1A* locus (8:7:0 for WT:heterozygous:homozygous KO), suggesting that *1a1b* may be lethal. The germination rates of seeds recovered from the mutant lines were comparable with WT plants (WT, 93%; 1A*1a*1B*1b*, 93%; and 1A*1a1b1b*, 96% of ~100 seeds), indicating that selection against a *1a1b* genotype occurred prior to seed development.

gRNA targeting strategy and transient expression of Cas9/gRNAs in protoplasts

To generate new *rbcs* mutants for individual members of the Arabidopsis SSU family using the CRISPR/Cas9 method, we initially designed two unique pairs of gRNAs to target specific regions of each of the four *rbcS* genes, for a total of eight gRNA pairs (Fig. 1A; Supplementary Table S1). Furthermore, to knock out *1B*, *2B*, and *3B* simultaneously, we designed a promiscuous pair of gRNAs to target homologous regions in the *1B–3B* locus (Fig. 1B; Supplementary Table S1). A paired gRNA approach was chosen for two reasons: (i) to increase the probability of generating mutations and (ii) the generation of larger deletions between the two gRNAs could potentially be screened more easily and cheaply (e.g. by PCR). Each gRNA pair was assembled using the Plant MoClo system as an individual expression cassette in a Level 2 binary vector containing a Cas9 expression cassette (Supplementary Data S1) (Engler *et al.*, 2014).

Vectors with each *rbcS*-specific gRNA pair were initially tested using a novel Arabidopsis protoplast transient expression system to estimate the efficiency of generating deletion events (Fig. 2A). Following transfection of the gRNA pairs targeting *1A* or *1B*, amplification of the respective gene loci produced the expected WT band and a second lower band indicating a deletion event based on the target sites of the gRNA pair. Pairs 1AP2 and 1BP2 most consistently produced prominent 'deletion bands' for *1A* and *1B*, respectively, and thus were selected for *in planta* expression. Only one gRNA pair each for *2B* (2BP2) and *3B* (3BP1) produced a deletion band. Sequence analysis of the deletion bands indicated cleavage 3–4 bp upstream of the gRNA PAM (photospacer adjacent motif) sites, consistent with the activity of SpCas9, and deletions ranging from 96 bp to 180 bp (Fig. 2B).

Expression of CRISPR/Cas9 in planta *to generate stable mutants for each* rbcS *isoform*

Arabidopsis plants were stably transformed with CRISPR/Cas9 binary vectors containing gRNA pairs 1AP2, 1BP2, 2BP2, 3BP1, and the promiscuous pair 1B3B. Transformed T_1 seeds were visually selected using the pFAST red fluorescent seed coat marker for the presence of Cas9 and gRNAs (Shimada *et al.*, 2010). The loci of the five gRNA pairs were initially screened by PCR for large deletions, which were detected in transformants for 1AP1, 2BP2, and 3BP1, but not for 1BP2 (Supplementary Fig. S5A). The deletion bands were consistent with those observed in protoplasts (Fig. 2A). Several deletion bands were detected for 1B3B that were consistent with predicted amplicon sizes following multiple cleavage events within the *1B–3B* locus (Supplementary Fig. S5B). Sequencing of each amplicon showed the expected cleavage position upstream of the PAM sites.

The efficiency with which deletions were induced by gRNA pairs ranged from 1% to 14% (Table 1). However, all deletions that were detected were accompanied by a significantly brighter WT band, indicating chimeric rather than heritable genomic mutations (i.e. homozygous or heterozygous) (Feng *et al.*, 2014; Pauwels *et al.*, 2018). T_1 transformants were also screened by Sanger sequencing to detect potential small indel mutations at each gRNA target site (Supplementary Fig. S5C). Indels were generally detected at a higher frequency than large deletions (3–31%).

Screening for heritable mutations in the T_2 generation was performed on segregated non-red seed progeny (i.e. containing no CRISPR/Cas9 insertion). Eight progeny from each T_1 line that showed large deletions were screened by PCR (Table 1). No heritable large deletions were detected for any gRNA pair, confirming that the observed deletions in T_1 were chimeric mutations. Sanger sequencing was then performed on T_2 lines that contained large deletions or indels in T_1. Stable homozygous indels (i.e. frameshift mutations in exonic regions that produced early stop codons) were identified for all four *rbcS* genes targeted by specific gRNA pairs (Supplementary Table S3; Supplementary Fig. S6). For the promiscuous gRNA pair 1B3B, each of the B subfamily genes was sequenced in 64 non-red T_2 plants from eight T_1 lines showing large deletions. No mutations were observed in *1B* in any of the lines tested. However, we did identify a 147 bp deletion in *3B* common to the progeny of a single T_1 line. Sequencing of the eight progeny revealed an additional 4 bp indel deletion in *2B* in a single plant, thus producing a *2b3b* double mutant. For *2b3b*, both indels led to frameshift mutations that produced early stop codons in *2B* and *3B*. For all *rbcs* lines, T_2 plants containing homozygous mutations were selected and seeds were collected for subsequent molecular and physiological characterizations in the T_3 generation.

Generating triple rbcs *mutants using the* 1a2b *T-DNA insertion mutant*

We next introduced CRISPR/Cas9 binary vectors containing gRNA pairs 3BP1 or 1BP2 into the new *1a2b* double T-DNA

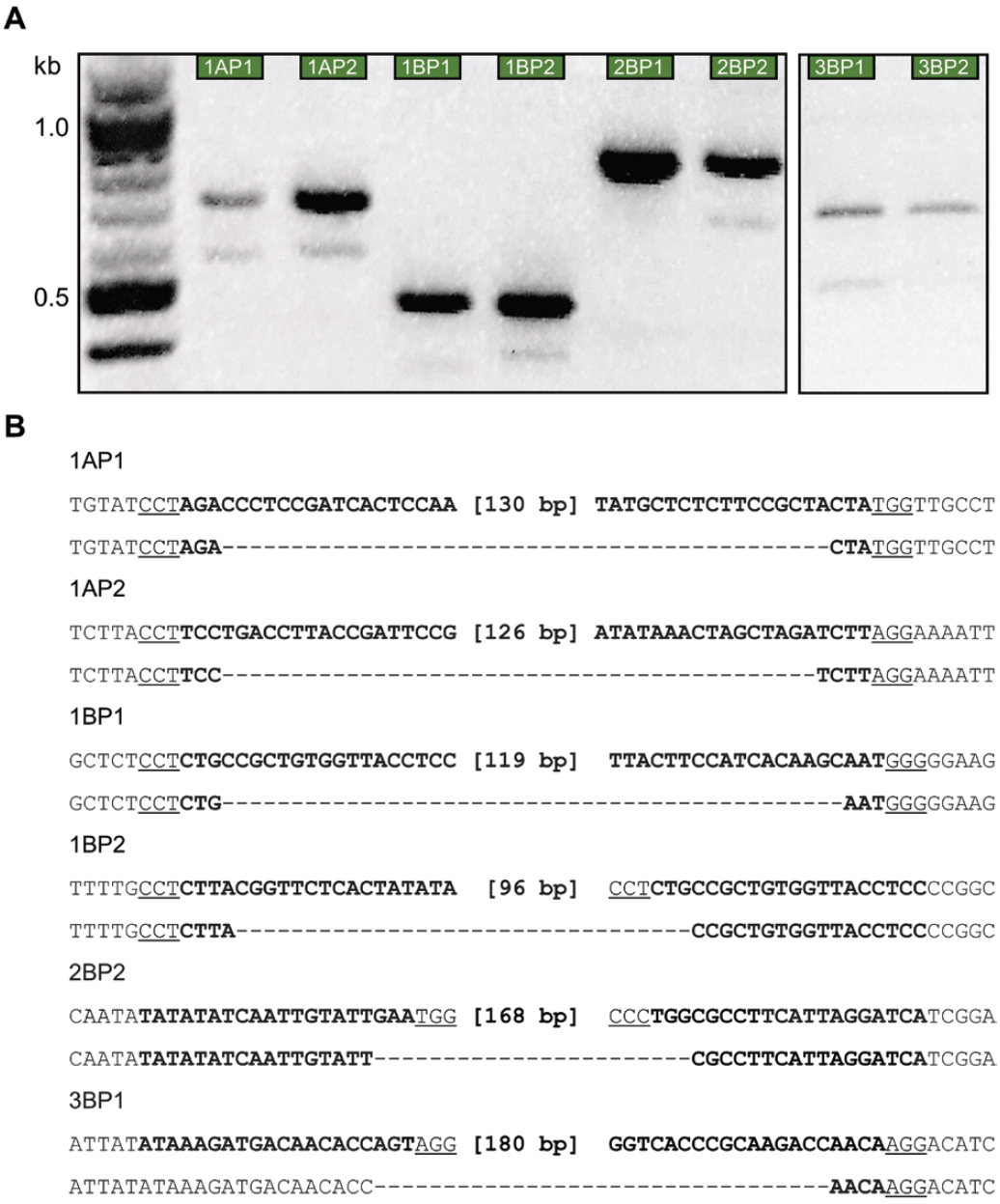

Fig. 2. Targeted mutagenesis of Rubisco small subunits in protoplasts of *Arabidopsis thaliana* using vectors encoding Cas9 and gRNA pairs. (A) Detection of mutations using PCR. Lane 1, DNA marker; lanes 2–9, PCR products of genomic DNA from protoplasts transfected with vectors carrying gRNA pairs targeting Rubisco small subunit (SSU) genes *1A* (1AP1, 1AP2), *1B* (1BP1, 1BP2), *2B* (2BP1, 2BP2), and *3B* (3BP1, 3BP2) (see Supplementary Table S2 for primer details). (B) Sequencing results of the lower 'deletion band' showed deletion events 3–4 bp upstream of the PAM site (underlined). The gRNA target sequences are shown in bold.

insertion mutant to attempt to produce the novel triple mutants *1a2b3b or 1a1b2b*. Out of 30 T_1 *1a2b* transformants for 3BP1, three showed a slow-growth, pale leaf phenotype compared with the *1a2b* phenotype (Table 1; Fig. 3A). Sanger sequencing of the *3B* locus confirmed the presence of a homozygous and biallelic frameshift mutation in all three T_1 plants (Supplementary Fig. S6). The heritability of those mutations was confirmed in non-red seed progeny of the T_2 generation for each line. In contrast, 1B is a minor SSU isoform, so no growth phenotype was expected in T_1 *1a2b* transformants for 1BP2. Sanger sequencing of the *1B* locus in 33 T_1 plants showed mutations in *1B* for eight plants (~ 25%), indicating that the CRISPR/Cas9 was functional and efficient. However, in all cases, only single base pair changes (i.e. a single codon substitution) or silent substitutions were observed (Supplementary Fig. S7).

Molecular characterisation of rbcs *mutants*

The expression profiles of the *rbcS* family and *rbcL* were quantified in T_3 plants for each *rbcS* mutant line. Transcript abundances were generally reduced for *rbcS* genes in lines targeted by specific T-DNA insertions or CRISPR/Cas9 editing (Supplementary Table S4). Consistent with previous observations, the relative expression of *rbcL* was more repressed in mutants with greater reductions in overall *rbcS* expression (i.e.

Table 1. Editing efficiency of paired gRNAs targeting each *rbcS* gene

gRNA pair	Target gene	Background	Number of T_1 plants (with transgene)		Number of T_2 plants (transgene free)	
			Large deletion	Indels/PM	Large deletion	Indels/PM
1AP2	1A	WT	8/92 (9%)	26/106 (25%)	0/8 (0%)	24/204 (12%)
1BP2	1B		0/112 (0%)	11/35 (31%)	N/A	42/66 (64%)
2BP2	2B		1/69 (1%)	12/63 (19%)	0/1 (0%)	30/78 (38%)
3BP1	3B		10/70 (14%)	1/32 (3%)	0/10 (0%)	24/66 (36%)
1B3B	1B	WT	8/76 (11%)	N/A	0/8 (0%)	0/64 (0%)
	2B					1/64 (2%)
	3B					8/64 (13%)
1BP2	1B	*1a2b*	0/33 (0%)	8/33 (24%)	N/A	N/A
3BP2	3B	*1a2b*	0/30 (0%)	3/30 (10%)	N/A	N/A

T_1 plants containing Cas9 were screened by PCR for large deletions and Sanger sequencing for indels and point mutations (PMs). T_1 lines containing large deletions or indels/PMs were screened for heritable mutations in Cas9-free plants in the T_2 generation. Six of the Cas9-free T_2 plants were sequenced for each T_1 line (34 lines for 1AP2, 11 lines for 1BP2, 13 for 2BP2, 11 lines for 3BP1, and 8 lines for 1B3B).

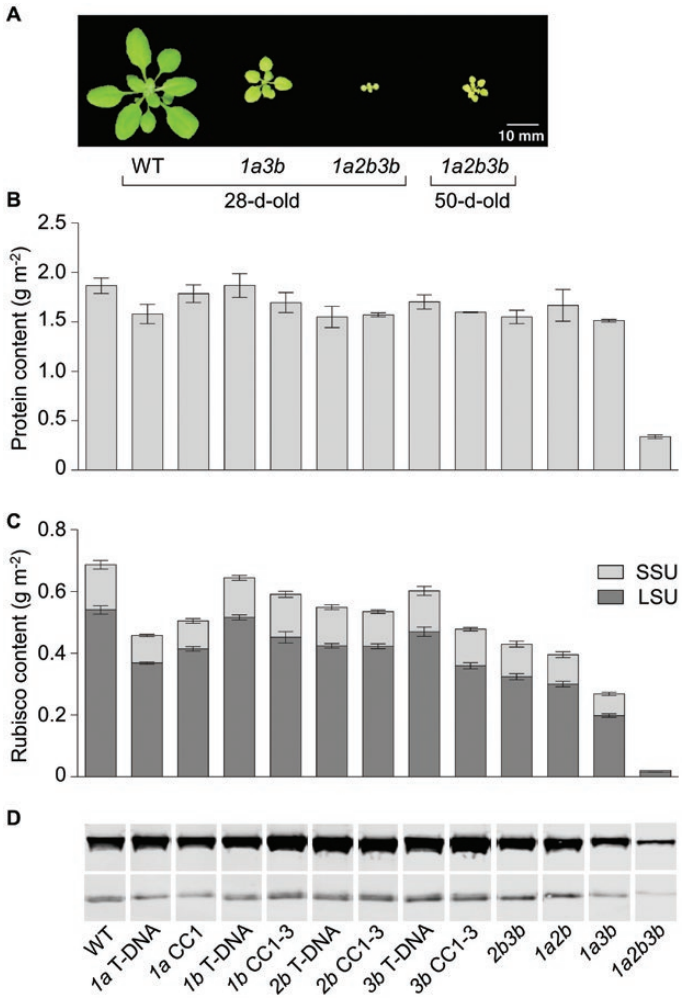

Fig. 3. Total soluble protein and Rubisco contents in *rbcs* mutants of *Arabidopsis thaliana*. (A) The *1a2b3b* mutant compared with *1a3b* and WT plants grown under standard conditions. All images represent 28-day-old rosettes unless otherwise stated. (B) Total soluble protein content shown for 35-day-old plants. (C) Rubisco content determined by [^{14}C]CABP binding and subunit ratios estimated by immunoblotting. Values are means ±SE of five measurements. (D) Representative immunoblots of *rbcs* mutants probed with a serum containing polyclonal antibodies against Rubisco to illustrate the reduction or absence of each SSU subfamily. The LSU (55 kDa) and SSU (14.7 kDa for 1A and 14.8 kDa for the B subfamily) are shown. See Supplementary Table S5 for further results and statistical analysis.

1a3b and *1a2b3b*) (Wostrikoff and Stern, 2007; Izumi *et al.*, 2012; Wostrikoff *et al.*, 2012; Atkinson *et al.*, 2017).

The total soluble protein content in leaves of all single mutants, and double mutants *1a2b* and *2b3b*, was similar to that of WT plants, but was significantly reduced by 20% and 82% for *1a3b* and *1a2b3b*, respectively (Fig. 3B; Supplementary Table S5). Leaf Rubisco content was generally decreased in single *rbcS* mutants compared with WT plants, and was significantly reduced in all double mutants and *1a2b3b* (Fig. 3C). Specifically, *1a2b* and *2b3b* had Rubisco levels reduced by 42% and 38%, respectively, *1a3b* by 61%, and *1a2b3b* by 97% relative to WT plants. Disruption of *1A* led to an absence of the lower 1A SSU (14.7 kDa) band in *1a*, *1a2b*, *1a3b*, and *1a2b3b* mutants as detected by western blot (Fig. 3D). Similarly, disruption of *3B* resulted in a reduction in the intensity of the upper SSU band of the B subfamily genes (14.8 kDa) in *3b*, *2b3b*, *1a3b*, and *1a2b3b* mutants. In contrast, no observable reductions in upper band intensity were observed with mutants of the less expressed isoforms *1B* and *2B*. Nevertheless, the SSU band intensity of *1a2b3b* was 30-fold lower than that of the overall intensity of the WT SSU bands.

Characterization of rbcs *mutants under normal light and high light*

Growth assays were performed under standard lab conditions for Arabidopsis plants (i.e. PPFD of 200 µmol photon m^{-2} s^{-1}) and under near-outdoor levels of light (i.e. PPFD of 1000 µmol photon m^{-2} s^{-1}) to characterize the growth performance of the *rbcs* mutants. Under standard lab conditions, the gene KO mutants for individual *rbcS* genes generated by CRISPR/Cas9 or T-DNA insertion (i.e. *1a*, *1b*, *2b*, and *3b*), as well as double mutants *1a2b* and *2b3b*, did not differ from WT plants in terms of rosette expansion rate, FW, DW, specific leaf area (SLA), maximum potential quantum efficiency of PSII (F_v/F_m), or chlorophyll content (Fig. 4; Supplementary Table S6). In contrast, *1a3b* and *1a2b3b* showed significant reductions in area, FW, DW, SLA, chlorophyll content, and F_v/F_m. *1a2b3b* was more slow growing than *1a3b*, but eventually developed to a similar level under standard lab conditions and produced viable seeds (Fig. 3A).

Under high light, *1a* and all three double mutants had significantly lower area, FW, DW, and chlorophyll content compared with the WT. A significant reduction in FW and DW was also observed for *3b* mutants, while rosette area remained similar to that of the WT. As a result, *3b* mutants also had a significantly higher SLA than the WT. High light was lethal to the *1a2b3b* triple mutant, which did not survive past 15 d after germination.

The response of A to C_i under saturating light (A/C_i curves) was measured for all *rbcs* mutant lines grown under standard lab conditions (Fig. 5) as well as key photosynthetic variables (Table 2). Stomatal conductance to CO_2 at ambient CO_2 (g_s) and respiration rates in the dark (R_d) were the same in all lines, consistent with previous work showing that reductions in Rubisco do not affect stomatal behaviour or mitochondrial respiration in Arabidopsis (Atkinson *et al.*, 2017). Single *rbcs* mutants for *1A* and *3B* showed significant decreases in the maximum rate of Rubisco carboxylation (V_{cmax}) compared with WT plants. In contrast, *1b* and *2b* mutants were similar to the WT. All three double mutants (*1a2b*, *2b3b*, and *1a3b*) had decreased V_{cmax}, and *1a3b* also showed an increased substomatal CO_2 compensation point (Γ). However, *1a2b3b* was significantly different from all other plant lines, with a 4-fold higher Γ value, and V_{cmax} at 10% of WT values.

Discussion

Although the uptake of CRISPR/Cas9 in plant biology has increased dramatically in recent years, many aspects concerning efficiency are still unclear, especially as established guidelines for CRISPR/Cas9 in other biological systems are not necessarily applicable to plants (Liang *et al.*, 2016; Hahn and Nekrasov, 2019). As the generation of stable plant transformants is time-consuming, we adopted a more rapid, transient expression approach using Arabidopsis protoplasts to test the efficiency of different gRNA pairs and select robust candidates (Fig. 2). All selected gRNA pairs were active in Arabidopsis and resulted in stable *rbcs* mutants. Thus, in agreement with previous studies (e.g. Li *et al.*, 2013; Durr *et al.*, 2018), screening gRNAs in protoplasts appears to be a reliable method for selecting functional gRNAs *in planta*. Overall, CRISPR/Cas appeared a viable approach to examine the impact of mutating *rbcS* isoforms and exploring their functional roles in plants.

CRISPR/Cas9 editing can be utilized to edit closely linked genes, which is not technically feasible through crossing T-DNA insertion lines, such as genes in the *rbcS* B subfamily. Here, the *2b3b* double *rbcs* mutant was successfully generated by targeting homologous regions by the gRNA pair 1B3B. In addition, the triple mutant *1a2b3b* was generated by targeting *3B* in the *1a2b* T-DNA line. Interestingly, disruption of *1B* in conjunction with other *rbcS* genes was not accomplished using T-DNA insertion or CRISPR/Cas9 approaches despite successful generation of *1b* mutants with both methods (Supplementary Fig. S4). 1B is the least expressed *rbcS* isoform and the mature peptide differs from 2B and 3B by only two amino acid residues (Izumi *et al.*, 2012). Thus, it is unlikely that the structure of 1B has a unique impact on Rubisco activity (Valegård *et al.*, 2018). Nevertheless, the temporal and/or spatial expression of 1B may contribute a specialized functional role that is critical for Arabidopsis development or fitness (i.e. during procreation). For example, previous work has shown that 1B localizes exclusively to the abaxial side of primordia and young leaves (Sawchuk *et al.*, 2008). It would be interesting to further examine 1B localization in reproductive tissues (e.g. siliques or flowers). As a product of gene duplication, *1B* is subject to selection pressure and is not retained in all accessions of Arabidopsis, although loss was correlated with disruption of the promoter (Schwarte and Tiedemann, 2011). Retention of 1B may be linked to sublocalization (Qiu *et al.*, 2019). Differential expression patterns of *rbcS* genes in different organs have also been observed in other plant species including tomato, maize, tobacco, and rice (Wanner and Gruissem, 1991; Ewing *et al.*, 1998; Morita, *et al.*, 2014; Laterre *et al.*, 2017). However, the extent of how important these organ-specific SSUs are has not yet been explored.

Gene editing via CRISPR/Cas9 resulted in a general decrease in mRNA abundance of target genes and significant reductions in Rubisco content for all *1a*, *1a2b*, *1a3b*, and *1a2b3b* mutants (Fig. 3; Supplementary Table S4). The observed reduction in mRNA levels was probably due to the presence of early stop codons that emerged from frameshift mutations, which consequently led to the degradation of mRNA through the nonsense-mediated decay process (Hug *et al.*, 2015). Reductions in *rbcS* transcripts led to a reduction in the *rbcL* transcript but not to the same extent, as Arabidopsis *rbcL* transcript is controlled post-transcriptionally at the translation initiation process (Rodermel *et al.*, 1996). Similar to Arabidopsis, *rbcL* synthesis in tobacco was shown to be partially independent of the *rbcS* transcript level, and LSU production was subject to state-dependent regulation of assembly that operated at the translational level (Wostrikoff and Stern, 2007).

rbcs mutants with a relatively small reduction in Rubisco content (i.e. <40% of WT levels) showed no change in growth rate and biomass accumulation relative to WT plants under standard lab growth conditions (i.e. PPFD of 200 μmol photons m^{-2} s^{-1}) (Fig. 4; Supplementary Table S6). This was not unexpected as reductions in Rubisco content have previously been shown to be compensated by an increase in activation state of the remaining Rubisco pool under conditions that are non-limiting for Rubisco (Quick *et al.*, 1991). However, under high light conditions (i.e. PPFD of 1000 μmol photons m^{-2} s^{-1}), biomass accumulation (i.e. FW and DW) was significantly decreased in *1a* and *3b* mutants and all three double mutants (i.e. *2b3b*, *1a2b*, and *1a3b*). This demonstrates that under near-outdoor levels of light, both 1A and 3B are critical for normal growth.

Differential expression of the Arabidopsis *rbcS* genes has been observed previously [e.g. under changing temperatures, light qualities, and CO_2 concentrations (Dedonder *et al.*, 1993; Cheng *et al.*, 1998; Yoon *et al.*, 2001; Sawchuk *et al.*, 2008)] but the impact of these differences on growth remains unclear. One potential use of our *rbcS* mutant suite could be to further examine the importance and interaction of specific SSU isoforms during development. All SSUs with the exception of *1B* are increased in response to specific wavelengths of light

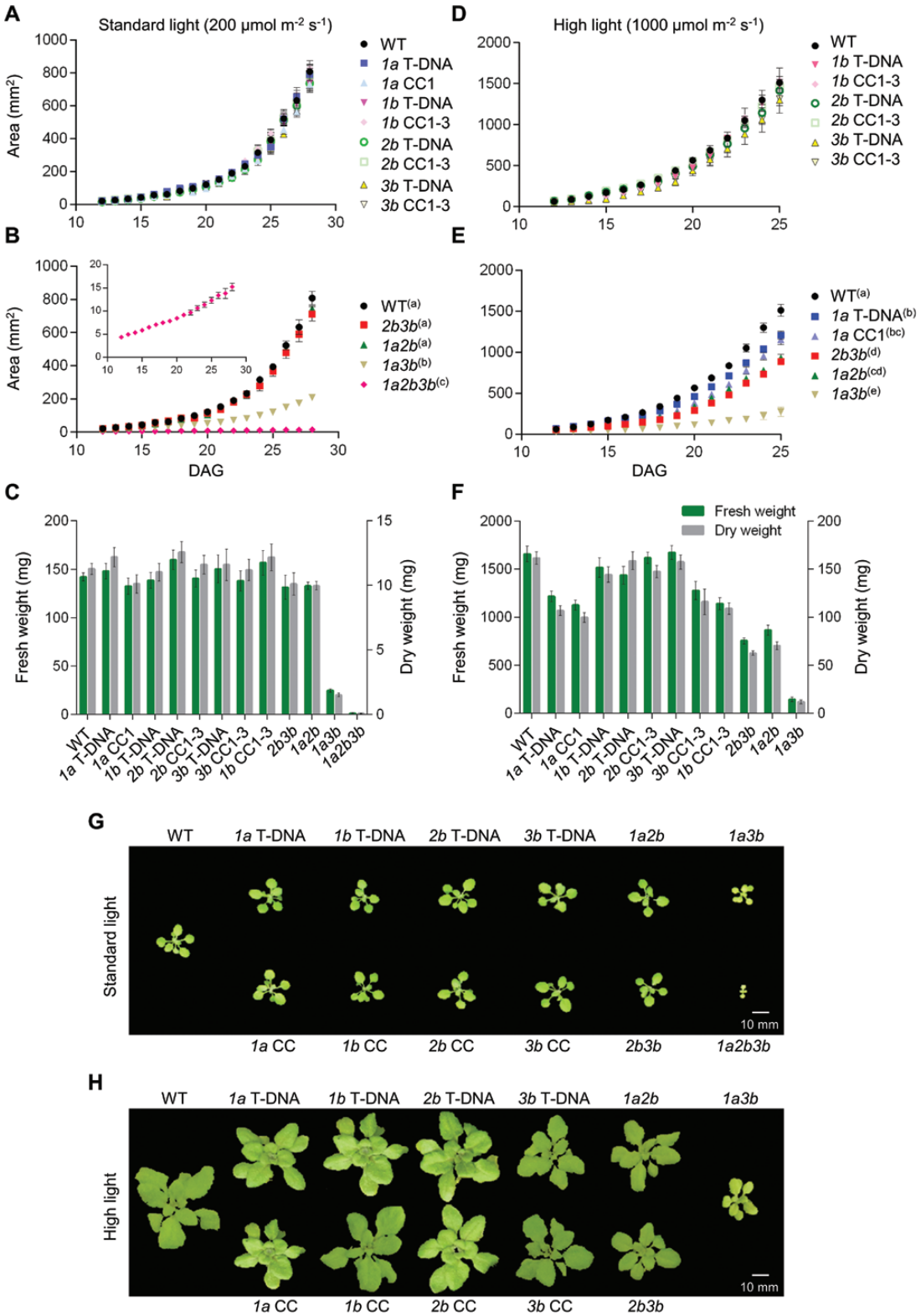

Fig. 4. Growth phenotypes of *rbcs* mutants of *Arabidopsis thaliana* grown under standard and high light conditions. (A) Rosette area expansion of *rbcs* single T-DNA and CRISPR/Cas9 (CC) mutants, (B) rosette area expansion of double and triple mutants (the insert shows the expansion of *1a2b3b* at appropriate scale), and (C) fresh and dry weights of 28-day-old rosettes of all mutant lines grown under standard conditions. (D) Rosette area expansion of *1b*, *2b*, and *3b* T-DNA and CRISPR/Cas9 mutants, (E) rosette area expansion of *1a*, *2b3b*, *1a2b*, and *1a3b* mutants, and (F) fresh and dry weights of 28-day-old rosettes for all mutant lines grown under high light conditions. (G) Representative examples of 20-day-old rosettes of plants grown in standard conditions and (H) under high light conditions. Letters in the keys for (B) and (E) indicate a significant difference between plant lines ($P < 0.05$) as determined by repeated measures ANOVA followed by Tukey's HSD tests. See Supplementary Table S6 for further results and statistical analysis.

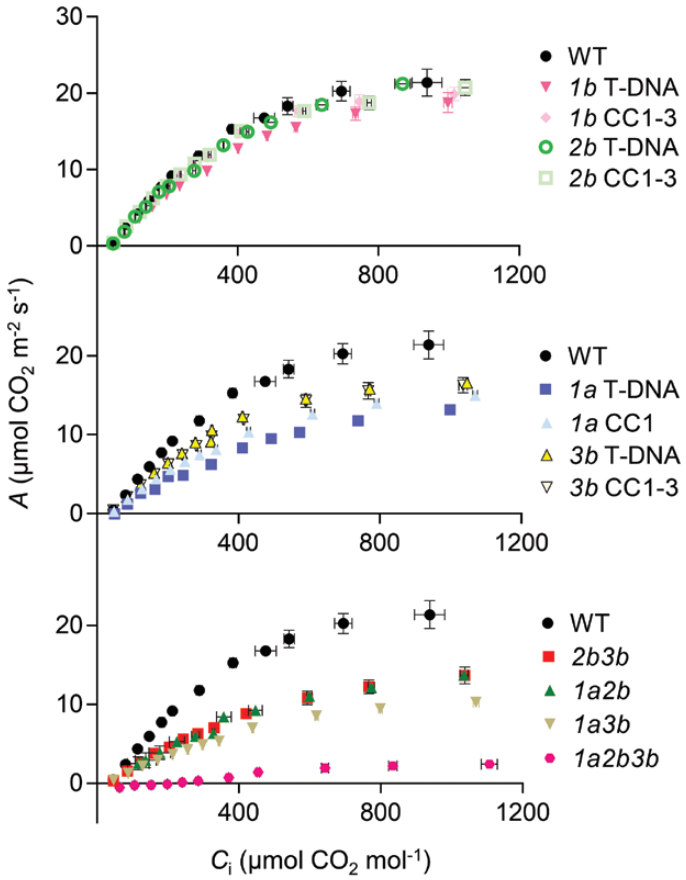

Fig. 5. Photosynthetic CO_2 response curves of *rbcs* mutants of *Arabidopsis thaliana*. Measurements were made on the fully expanded sixth or seventh leaf of 35- to 45-day-old non-flowering rosettes for the WT and mutants and 80-day-old non-flowering rosettes for *1a2b3b*. The A/C_i curves show the response of net CO_2 assimilation (A) to different substomatal concentration of CO_2 (C_i) under saturating light (1800 µmol photon m^{-2} s^{-1}) for (A) *1b* and *2b* mutants, (B) *1a* and *3b* mutants, and (C) *2b3b*, *1a2b*, *1a3b*, and *1a2b3b* mutants. Each value represents the means ±SE of measurements made on individual leaves from 3–4 different rosettes.

(i.e. blue, red, and far-red light), but at differing relative gene expression levels (Dedonder *et al.*, 1993; Sawchuk *et al.*, 2008). Furthermore, *1A* is the major isoform below 20 °C, whereas *3B* expression is dominant at 30 °C (Yoon *et al.*, 2001; Izumi *et al.*, 2012).

The rosette area of *3b* mutants was not decreased under high light compared with WT plants (Fig. 4; Supplementary Table S6). However, the observed reduction in biomass resulted in a significant increase in SLA and suggested that leaves were thinner in those lines. Furthermore, gas exchange measurements demonstrated a decrease in V_{cmax} for *3b* mutants (Fig. 5; Table 2). Similar increases in SLA and reductions in photosynthetic capacity were observed for *1a* and all three double mutants. These observations are in line with previous growth analyses of Arabidopsis Rubisco activase antisense lines under high light (Eckardt *et al.*, 1997), where a reallocation of resources to expand leaf area and reduce thickness was observed when photosynthetic capacity was limiting (Hoshino *et al.*, 2019).

For the most severely Rubisco-limited mutants, *1a3b* and *1a2b3b*, F_v/F_m was reduced under standard light (Supplementary Table S6). However, F_v/F_m for *1a3b* was similar to that of the WT under high light, indicating that a reduction in Rubisco has less impact on the operating efficiency of the light reactions under high light. Thus, failure of *1a2b3b* to grow under high light could indicate an inability of the light reactions to coordinate product utilization (i.e. ATP and NADPH) with the extremely low Rubisco content of the triple mutant.

CRISPR/Cas9 is a versatile tool that has been successfully used for genetic editing and the enhancement of breeding strategies in a wide variety of plant and crop species (Khumsupan *et al.*, 2019; Wolter *et al.*, 2019). This study has shown that CRISPR/Cas is a viable approach for characterizing the roles of SSUs in plant species and that Arabidopsis mutants lacking SSU isoforms are useful platforms for the study of functional roles of SSUs. In particular, the triple mutant *1a2b3b* is potentially a powerful resource for studying the impact of

Table 2. Variables derived from photosynthetic CO_2 response curves, based on leaf gas exchange analysis

	V_{cmax} (µmol CO_2 m^{-2} s^{-1})	R_d (µmol CO_2 m^{-2} s^{-1})	g_s (mmol H_2O m^{-2} s^{-1})	Γ (µmol CO_2 mol^{-1})
WT	55.5 ± 3.5 a	1.01 ± 0.27 a	0.25 ± 0.03 a	52.4 ± 4.7 a
1a T-DNA	38.0 ± 3.2 bc	1.18 ± 0.11 a	0.20 ± 0.03 a	65.7 ± 10.1 ab
1a CC1	38.1 ± 2.0 bc	0.93 ± 0.08 a	0.25 ± 0.01 a	61.3 ± 1.0 ab
1b T-DNA	50.3 ± 5.2 ab	1.20 ± 0.11 a	0.24 ± 0.01 a	63.2 ± 3.2 ab
1b CC1-3	49.4 ± 1.4 ab	0.83 ± 0.08 a	0.28 ± 0.01 a	53.8 ± 1.3 a
2b T-DNA	48.2 ± 4.7 ab	1.17 ± 0.08 a	0.24 ± 0.07 a	60.1 ± 2.3 ab
2b CC1-3	48.9 ± 1.5 ab	0.81 ± 0.06 a	0.28 ± 0.02 a	52.7 ± 1.3 a
3b T-DNA	39.8 ± 2.0 b	0.73 ± 0.05 a	0.26 ± 0.01 a	52.5 ± 0.8 a
3b CC1-3	38.5 ± 1.5 bc	0.77 ± 0.03 a	0.27 ± 0.01 a	54.7 ± 2.4 a
2b3b	36.9 ± 3.2 bc	0.85 ± 0.07 a	0.25 ± 0.02 a	65.3 ± 2.9 ab
1a2b	36.3 ± 2.9 bc	0.78 ± 0.08 a	0.27 ± 0.01 a	64.7 ± 2.3 ab
1a3b	24.6 ± 1.0 c	0.71 ± 0.04 a	0.26 ± 0.01 a	70.0 ± 1.5 b
1a2b3b	5.4 ± 0.3 d	0.84 ± 0.13 a	0.23 ± 0.04 a	200.1 ± 3.8 c

Values are means ±SE of measurements made on three or four leaves from different plants (35- to 45-day-old non-flowering rosettes for the WT and mutants and 80-day-old non-flowering rosettes for *1a2b3b*). Values followed by the same letters in the same column are not significantly different ($P < 0.05$) as determined by ANOVA followed by Tukey's HSD tests. Abbreviations: Γ, substomatal CO_2 compensation point; g_s, stomatal conductance at 400 ppm CO_2; R_d, mitochondrial respiration in the light; V_{cmax}, maximum rate of Rubisco carboxylation.

heterologous SSU expression that has previously been studied in the *1a3b* background (Atkinson *et al.*, 2017). Unlike *1a3b*, which has a Rubisco content of ~35% relative to the WT (Izumi *et al.*, 2012; Atkinson *et al.*, 2017), *1a2b3b* retained only 3% under the conditions tested (Fig. 3; Supplementary Table S5), and is the first example of a plant line with a homogenous Rubisco pool consisting of single SSU and a single LSU isoform. Following complementation, further disruption of the remaining *1B* isoform (e.g. by CRISPR/Cas9 editing using the gRNA pair 1BP2) could be used to generate a true hybrid Rubisco pool comprised of only heterologous SSU(s) and the native LSU. In addition, as the contribution of individual SSUs to the Rubisco enzyme is still unclear, the triple mutant could be exploited as a model to knock in native SSUs tagged with different fluorescent probes (Ishida *et al.*, 2008). This method would allow for the visualization of composition of each hexadecamer in the Rubisco enzyme.

Supplementary data

Supplementary data are available at *JXB* online.

Table S1. List of CRISPR/Cas9 vectors constructed in this study.

Table S2. Sequences of synthetic oligonucleotides used in this study.

Table S3. Editing efficiency of individual gRNAs for each pair of gRNAs.

Table S4. Transcript abundances of the Rubisco gene family in *rbcs* mutants of *Arabidopsis thaliana*.

Table S5. Rubisco and total soluble protein contents for *rbcs* mutants of *Arabidopsis thaliana*.

Table S6. Rosette area and biomass of *rbcs* mutants of *Arabidopsis thaliana* grown at standard versus high light conditions.

Fig. S1. Synthesized nucleic acid sequence containing 3'-UTRs of *rbcS* genes used for RT–qPCR quantification.

Fig. S2. Light response curve of Arabidopsis WT plants.

Fig. S3. Segregation of a new *1a2b* double T-DNA insertion mutant.

Fig. S4. Segregation ratios for progeny of *1a* and *1b* T-DNA insertion mutants of *Arabidopsis thaliana*.

Fig. S5. Screening for large deletion and indel events in *rbcs* mutants of *Arabidopsis thaliana*.

Fig. S6. A library of *Arabidopsis thaliana rbcs* mutants produced using CRISPR/Cas9.

Fig. S7. Sequence alignments of the target site for the 1BP2 CRISPR/Cas9 construct in *1a2b* T-DNA mutants.

Data S1. Zip file containing sequence maps (.gb files) of the vectors used in this study.

Acknowledgements

This work was supported by the UK Biotechnology and Biological Sciences Research Council (grants BB/I024488/1 to EC-S, and BB/M006468/1 and BB/S015531/1 to AJM) and the Leverhulme Trust (grant RPG-2017–402 to AJM). PK was funded by a postgraduate research scholarship from the Darwin Trust of Edinburgh. We thank Louis Caruana (Lancaster University) for technical support.

References

Andersson I, Backlund A. 2008. Structure and function of Rubisco. Plant Physiology and Biochemistry **46**, 275–291.

Atkinson N, Leitão N, Orr DJ, Meyer MT, Carmo-Silva E, Griffiths H, Smith AM, McCormick AJ. 2017. Rubisco small subunits from the unicellular green alga Chlamydomonas complement Rubisco-deficient mutants of Arabidopsis. New Phytologist **214**, 655–667.

Bellasio C, Beerling DJ, Griffiths H. 2016. An excel tool for deriving key photosynthetic parameters from combined gas exchange and chlorophyll fluorescence: theory and practice. Plant, Cell & Environment **39**, 1180–1197.

Bracher A, Whitney SM, Hartl FU, Hayer-Hartl M. 2017. Biogenesis and metabolic maintenance of Rubisco. Annual Review of Plant Biology **68**, 29–60.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. Analytical Biochemistry **72**, 248–254.

Cavanagh AP, Kubien DS. 2014. Can phenotypic plasticity in Rubisco performance contribute to photosynthetic acclimation? Photosynthesis Research **119**, 203–214.

Cheng SH, Moore B, Seemann JR. 1998. Effects of short- and long-term elevated CO_2 on the expression of ribulose-1,5-bisphosphate carboxylase/oxygenase genes and carbohydrate accumulation in leaves of *Arabidopsis thaliana* (L.) Heynh. Plant Physiology **116**, 715–723.

Chupeau MC, Granier F, Pichon O, Renou JP, Gaudin V, Chupeau Y. 2013. Characterization of the early events leading to totipotency in an Arabidopsis protoplast liquid culture by temporal transcript profiling. The Plant Cell **25**, 2444–2463.

Clark K, Krysan P. 2010. Chromosomal translocations are a common phenomenon in *Arabidopsis thaliana* T-DNA insertion lines. The Plant Journal **6**, 990–1001.

Clough SJ, Bent AF. 1999. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. The Plant Journal **16**, 735–743.

Day CD, Lee E, Kobayashi J, Holappa LD, Albert H, Ow DW. 2000. Transgene integration into the same chromosome location can produce alleles that express at a predictable level or alleles that are differentially silenced. Genes and Development **14**, 2869–2880.

Dedonder A, Rethy R, Fredericq H, Van Montagu M, Krebbers E. 1993. Arabidopsis *RbcS* genes are differentially regulated by light. Plant Physiology **101**, 801–808.

Dobrescu A, Scorza LCT, Tsaftaris SA, McCormick AJ. 2017. A 'Do-It-Yourself' phenotyping system: measuring growth and morphology throughout the diel cycle in rosette shaped plants. Plant Methods **13**, 95.

Durr J, Papareddy R, Nakajima K, Gutierrez-Marcos J. 2018. Highly efficient heritable targeted deletions of gene clusters and non-coding regulatory regions in Arabidopsis using CRISPR/Cas9. Scientific Reports **8**, 4443.

Eckardt NA, Snyder GW, Portis AR Jr, Orgen WL. 1997. Growth and photosynthesis under high and low irradiance of *Arabidopsis thaliana* antisense mutants with reduced ribulose-1,5-bisphosphate carboxylase/oxygenase activase content. Plant Physiology **113**, 575–586.

Engler C, Youles M, Gruetzner R, Ehnert TM, Werner S, Jones JD, Patron NJ, Marillonnet S. 2014. A golden gate modular cloning toolbox for plants. ACS Synthetic Biology **3**, 839–843.

Esquivel MG, Genkov T, Nogueira AS, Salvucci ME, Spreitzer RJ. 2013. Substitutions at the opening of the Rubisco central solvent channel affect holoenzyme stability and CO_2/O_2 specificity but not activation by Rubisco activase. Photosynthesis Research **118**, 209–218.

Ethier GJ, Livingston NJ. 2004. On the need to incorporate sensitivity to CO_2 transfer conductance into the Farquhar–von Caemmerer–Berry leaf photosynthesis model. Plant, Cell & Environment **27**, 137–153.

Ewing RM, Jenkins GI, Langdale JA. 1998. Transcripts of maize *RbcS* genes accumulate differentially in C3 and C4 tissues. Plant Molecular Biology **36**, 593–599.

Feng Z, Mao Y, Xu N, *et al.* 2014. Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. Proceedings of the National Academy of Sciences, USA **111**, 4632–4637.

Fukayama H, Kobara T, Shiomi K, Morita R, Sasayama D, Hatanaka T, Azuma T. 2019. Rubisco small subunits of C4 plants, Napier grass and guinea grass confer C4-like catalytic properties on Rubisco in rice. Plant Production Science **22**, 296–300.

Galili S, Galili G, Avivi Y, Feldman M. 1992. Identification and chromosomal location of four subfamilies of the Rubisco small subunit genes in common wheat. Theoretical and Applied Genetics **83**, 385–391.

Genkov T, Spreitzer RJ. 2009. Highly conserved small subunit residues influence rubisco large subunit catalysis. Journal of Biological Chemistry **284**, 30105–30112.

Getzoff TP, Zhu G, Bohnert HJ, Jensen RG. 1998. Carboxylase/oxygenase containing a pea small subunit protein is compromised in carbamylation. Plant Physiology **116**, 695–702.

Hahn F, Nekrasov V. 2019. CRISPR/Cas precision: do we need to worry about off-targeting in plants? Plant Cell Reports **38**, 437–441.

Hoshino R, Yoshida Y, Tsukay H. 2019. Multiple steps of leaf thickening during sun-leaf formation in Arabidopsis. The Plant Journal **100**, 738–753.

Howe CJ, Auffret AD, Doherty A, Bowman CM, Dyer TA, Gray JC. 1982. Location and nucleotide sequence of the gene for the proton-translocating subunit of wheat chloroplast ATP synthase. Proceedings of the National Academy of Sciences, USA **79**, 6903–6907.

Hug N, Longman D, Cáceres JF. 2015. Mechanism and regulation of the nonsense-mediated decay pathway. Nucleic Acids Research **44**, 1483–1495.

Ishida H, Yoshimoto K, Izumi M, Reisen D, Yano Y, Makino A, Ohsumi Y, Hanson MR, Mae T. 2008. Mobilization of rubisco and stroma-localized fluorescent proteins of chloroplasts to the vacuole by an ATG gene-dependent autophagic process. Plant Physiology **148**, 142–155.

Ishikawa C, Hatanaka T, Misoo S, Miyake C, Fukayama H. 2011. Functional incorporation of sorghum small subunit increases the catalytic turnover rate of Rubisco in transgenic rice. Plant Physiology **156**, 1603–1611.

Izumi M, Tsunoda H, Suzuki Y, Makino A, Ishida H. 2012. RBCS1A and RBCS3B, two major members within the Arabidopsis RBCS multigene family, function to yield sufficient Rubisco content for leaf photosynthetic capacity. Journal of Experimental Botany **63**, 2159–2170.

Khumsupan P, Donovan S, McCormick AJ. 2019. CRISPR/Cas in Arabidopsis: overcoming challenges to accelerate improvements in crop photosynthetic efficiencies. Physiologia Plantarum **166**, 428–437.

Klepikova AV, Kasianov AS, Gerasimov ES, Logacheva MD, Penin AA. 2016. A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. The Plant Journal **88**, 1058–1070.

Kostov RV, Small CL, McFadden BA. 1997. Mutations in a sequence near the N-terminus of the small subunit alter the CO_2/O_2 specificity factor for ribulose bisphosphate carboxylase/oxygenase. Photosynthesis Resesearch **54**, 127–134.

Krebbers E, Seurinck J, Herdies L, Cashmore AR, Timko MP. 1988. Four genes in two diverged subfamilies encode the ribulose-1,5-bisphosphate carboxylase small subunit polypeptides of *Arabidopsis thaliana*. Plant Molecular Biology **11**, 745–759.

Kubis A, Bar-Even A. 2019. Synthetic biology approaches for improving photosynthesis. Journal of Experimental Botany **70**, 1425–1433.

Laterre R, Pottier M, Remacle C, Boutry M. 2017. Photosynthetic trichomes contain a specific Rubisco with a modified pH-dependent activity. Plant Physiology **173**, 2110–2120.

Li J, Chory J. 1998. Preparation of DNA from Arabidopsis. Methods in Molecular Biology **82**, 55–60.

Li W, Teng F, Li T, Zhou Q. 2013. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR–Cas systems. Nature Biotechnology **31**, 684–686.

Liang G, Zhang H, Lou D, Yu D. 2016. Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. Scientific Reports **19**, 1–8.

Lin MT, Stone WD, Chaudhari V, Hanson MR. 2019. Enzyme kinetics of tobacco Rubisco expressed in *Escherichia coli* varies depending on the small subunit composition. BioRxiv doi: 10.1101/562223. [Preprint].

Maxwell K, Johnson GN. 2000. Chlorophyll fluorescence—a practical guide. Journal of Experimental Botany **51**, 659–668.

McCormick AJ, Kruger NJ. 2015. Lack of fructose 2,6-bisphosphate compromises photosynthesis and growth in Arabidopsis in fluctuating environments. The Plant Journal **81**, 670–683.

Meier I, Callan KL, Fleming AJ, Gruissem W. 1995. Organ-specific differential regulation of a promoter subfamily for the ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit genes in tomato. Plant Physiology **107**, 1105–1118.

Morita K, Hatanaka T, Misoo S, Fukayama H. 2014. Unusual small subunit that is not expressed in photosynthetic cells alters the catalytic properties of rubisco in rice. Plant Physiology **164**, 69–79.

Nacry P, Camilleri C, Courtial B, Caboche M, Bouchez D. 1998. Major chromosomal rearrangements induced by T-DNA transformation in Arabidopsis. Genetics **149**, 641–650.

Nagaya S, Kawamura K, Shinmyo A, Kato K. 2010. The HSP terminator of *Arabidopsis thaliana* increases gene expression in plant cells. Plant & Cell Physiology **51**, 328–332.

Niwa Y, Goto K, Shimizu M, Kobayashi H. 1997. Chromosomal mapping of genes in the *RBCS* family in *Arabidopsis thaliana*. DNA Research **4**, 341–343.

Ogawa S, Suzuki Y, Yoshizawa R, Kanno K, Makino A. 2012. Effect of individual suppression of *RBCS* multigene family on Rubisco contents in rice leaves. Plant, Cell & Environment **35**, 546–553.

Orr DJ, Worrall D, Lin MT, Carmo-Silva E, Hanson MR, Parry MAJ. 2020. Hybrid cyanobacterial–tobacco Rubisco supports autotrophic growth and pro-carboxysomal aggregation. Plant Physiology **182**, 807–818.

Parry G, Patron N, Bastow R, Matthewman C. 2016. Meeting report: GARNet/OpenPlant CRISPR–Cas workshop. Plant Methods **12**, 6.

Patron N, Orzaez D, Marillonnet S, *et al.* 2015. Standards for plant synthetic biology: a common syntax for exchange of DNA parts. New Phytologist **208**, 13–19.

Paul K, Morell MK, Andrews TJ. 1991. Mutations in the small subunit of ribulosebisphosphate carboxylase affect subunit binding and catalysis. Biochemistry **30**, 10019–10026.

Pauwels L, De Clercq R, Goossens J, Iñigo S, Williams C, Ron M, Britt A, Goossens A. 2018. A dual sgRNA approach for functional genomics in *Arabidopsis thaliana*. G3 **8**, 2603–2615.

Porra RJ, Thompson WA, Kriedemann PE. 1989. Determination of accurate extinction coefficients and simultaneous-equations for assaying chlorophyll-a and chlorophyll-b extracted with 4 different solvents—verification of the concentration of chlorophyll standards by atomic-absorption spectroscopy. Biochimica et Biophysica Acta **975**, 384–394.

Pottier M, Gilis D, Boutry M. 2018. The hidden face of Rubisco. Trends in Plant Science **23**, 382–392.

Qiu Y, Tay Y, Van Ruan Y, Adams KL. 2019. Divergence of duplicated genes by repeated partitioning of splice forms and subcellular localization. New Phytologist **225**, 1011–1022.

Quick WP, Schurr U, Scheibe R, Schulze ED, Rodermel SR, Bogorad L, Stitt M. 1991. Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with 'antisense' *rbcS*: I. Impact on photosynthesis in ambient growth conditions. Planta **183**, 542–554.

Rae BD, Long BM, Förster B, Nguyen ND, Velanis CN, Atkinson N, Hee WY, Mukherjee B, Price GD, McCormick AJ. 2017. Progress and challenges of engineering a biophysical CO_2-concentrating mechanism into higher plants. Journal of Experimental Botany **68**, 3717–3737.

Raitskin O, Schudoma C, West A, Patron NJ. 2019. Comparison of efficiency and specificity of CRISPR-associated (Cas) nucleases in plants: an expanded toolkit for precision genome engineering. PLoS One **14**, e0211598.

Read BA, Tabita FR. 1992. A hybrid ribulosebisphosphate carboxylase/oxygenase enzyme exhibiting a substantial increase in substrate specificity factor. Biochemistry **31**, 5553–5560.

Rodermel S. 1999. Subunit control of Rubisco biosynthesis—a relic of an endosymbiotic past. Photosynthesis Research **59**, 105–123.

Rodermel S, Haley J, Jiang CZ, Tsai CH, Bogorad L. 1996. A mechanism for intergenomic integration: abundance of ribulose bisphosphate carboxylase small-subunit protein influences the translation of the

large-subunit mRNA. Proceedings of the National Academy of Sciences, USA **93**, 3881–3885.

Sasanuma T. 2001. Characterization of the *rbcS* multigene family in wheat: subfamily classification, determination of chromosomal location and evolutionary analysis. Molecular Genetics and Genomics **265**, 161–171.

Sawchuk MG, Donner TJ, Head P, Scarpella E. 2008. Unique and overlapping expression patterns among members of photosynthesis-associated nuclear gene families in Arabidopsis. Plant Physiology **148**, 1908–1924.

Schneider G, Knight S, Andersson I, Brändén CI, Lindqvist Y, Lundqvist T. 1990. Comparison of the crystal structures of L2 and L8S8 Rubisco suggests a functional role for the small subunit. The EMBO Journal **9**, 2045–2050.

Schwarte S, Tiedemann R. 2011. A gene duplication/loss event in the ribulose-1,5-bisphosphate-carboxylase/oxygenase (rubisco) small subunit gene family among accessions of *Arabidopsis thaliana*. Molecular Biology and Evolution **28**, 1861–1876.

Sharwood RE, Sonawane BV, Ghannoum O, Whitney SM. 2016. Improved analysis of C4 and C3 photosynthesis via refined in vitro assays of their carbon fixation biochemistry. Journal of Experimental Botany **67**, 3137–3148.

Shimada TL, Shimada T, Hara-Nishimura I. 2010. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. The Plant Journal **61**, 519–528.

Simkin AJ, López-Calcagno PE, Raines CA. 2019. Feeding the world: improving photosynthetic efficiency for sustainable crop production. Journal of Experimental Botany **70**, 1119–1140.

South PF, Cavanagh AP, Liu HW, Ort DR. 2019. Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. Science **363**, 1–10.

Spreitzer RJ, Peddi SR, Satagopan S. 2005. Phylogenetic engineering at an interface between large and small subunits imparts land-plant kinetic properties to algal Rubisco. Proceedings of the National Academy of Sciences, USA **102**, 17225–30.

Suzuki Y, Nakabayashi K, Yoshizawa R, Mae T, Makino A. 2009. Differences in expression of the *RBCS* multigene family and rubisco protein content in various rice plant tissues at different growth stages. Plant & Cell Physiology **50**, 1851–1855.

Valegård K, Hasse D, Andersson I, Gunn LH. 2018. Structure of Rubisco from *Arabidopsis thaliana* in complex with 2-carboxyarabinitol-1,5-bisphosphate. Acta Crystallographica. Section D, Structural Biology **74**, 1–9.

van Lun M, van der Spoel D, Andersson I. 2011. Subunit interface dynamics in hexadecameric rubisco. Journal of Molecular Biology **411**, 1083–1098.

Vasudevan R, Gale GAR, Schiavon AA, *et al.* 2019. CyanoGate: a modular cloning suite for engineering cyanobacteria based on the plant MoClo syntax. Plant Physiology **180**, 39–55.

Wanner LA, Gruissem W. 1991. Expression dynamics of the tomato *rbcS* gene family during development. The Plant Cell **3**, 1289–1303.

Wehmeyer B, Cashmore AR, Schäfer E. 1990. Photocontrol of the expression of genes encoding chlorophyll a/b binding proteins and small subunit of ribulose-1,5-bisphosphate carboxylase in etiolated seedlings of *Lycopersicon esculentum* (L.) and *Nicotiana tabacum* (L.). Plant Physiology **93**, 990–997.

Whitney SM, Birch R, Kelso C, Beck JL, Kapralov MV. 2015. Improving recombinant Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1 chaperone. Proceedings of the National Academy of Sciences, USA **112**, 3564–3569.

Wolter F, Schindele P, Puchta H. 2019. Plant breeding at the speed of light: the power of CRISPR/Cas to generate directed genetic diversity at multiple sites. BMC Plant Biology **19**, 176.

Wostrikoff K, Clark A, Sato S, Clemente T, Stern D. 2012. Ectopic expression of Rubisco subunits in maize mesophyll cells does not overcome barriers to cell type-specific accumulation. Plant Physiology **160**, 419–432.

Wostrikoff K, Stern D. 2007. Rubisco large-subunit translation is autoregulated in response to its assembly state in tobacco chloroplasts. Proceedings of the National Academy of Sciences, USA **104**, 6466–6471.

Yamada K, Davydov II, Besnard G, Salamin N. 2019. Duplication history and molecular evolution of the *rbcS* multigene family in angiosperms. Journal of Experimental Botany **70**, 6127–6139.

Yoon M, Putterill JJ, Ross GS, Laing WA. 2001. Determination of the relative expression levels of rubisco small subunit genes in Arabidopsis by rapid amplification of cDNA ends. Analytical Biochemistry **291**, 237–244.